

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
29 July 2004 (29.07.2004)

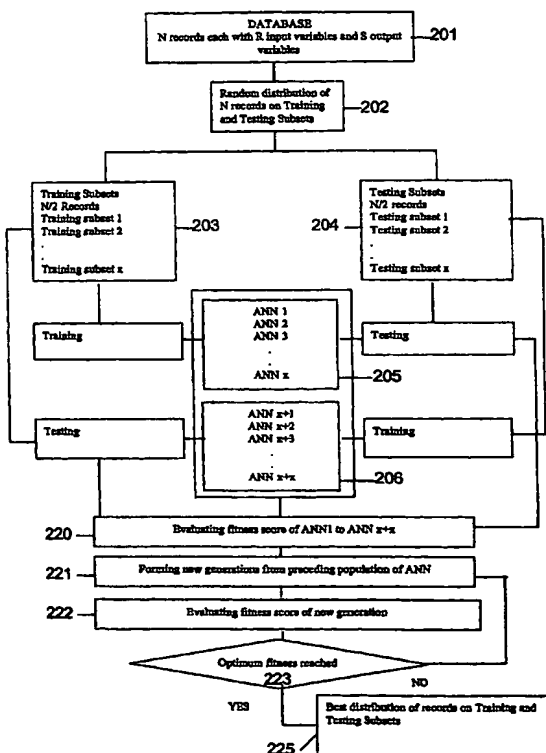
PCT

(10) International Publication Number  
WO 2004/063831 A3

- (51) International Patent Classification<sup>7</sup>: G06N 3/08, G06F 19/00
- (21) International Application Number: PCT/EP2004/000157
- (22) International Filing Date: 13 January 2004 (13.01.2004)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/440,210 15 January 2003 (15.01.2003) US
- (71) Applicant (for all designated States except US): BRACCO IMAGING S.P.A. [IT/IT]; 50 v. Folli, I-20134 Milano (IT).
- (72) Inventor; and
- (75) Inventor/Applicant (for US only): BUSCEMA, Massimo [IT/IT]; 47 v. N. Coviello, I-00163 Roma (IT).
- (74) Agent: KARAGHIOSOFF, Giorgio, A.; 27 V. Pecorile, I-17015 Celle Ligure (SV) (IT).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK,

[Continued on next page]

(54) Title: SYSTEM AND METHOD FOR OPTIMIZATION OF A DATABASE FOR THE TRAINING AND TESTING OF PREDICTION ALGORITHMS



(57) Abstract: A system and method are provided for the training and testing of prediction algorithms. According to an exemplary embodiment of the invention the method generates optimum training, testing and/or validation data sets from a common general database by applying a genetic algorithm to populations of testing and training subsets used in connection with a given prediction algorithm. In exemplary embodiments the prediction algorithm operated upon is an artificial neural network. As well, in preferred exemplary embodiments, the most predictive independent variables of the records of the common database are automatically selected in a pre-processing phase. Such preprocessing phase applies a genetic algorithm to populations of prediction algorithms which vary as to number and content of input variables, where the prediction algorithms representing the selections of input variables which have the best testing performances and the minimum input variables are promoted for the processing of the new generations according to a defined selection algorithm.



TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**Published:**

- *with international search report*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

**(88) Date of publication of the international search report:**  
11 August 2005

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/EP2004/000157

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 7 G06N3/08 G06F19/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 7 G06N G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, IBM-TDB, INSPEC

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y A	US 2002/059154 A1 (RODVOLD DAVID M) 16 May 2002 (2002-05-16) paragraph '0007! - paragraph '0053!	1,4, 23-25 30-34 2,3, 5-22, 26-29
Y	US 2002/184569 A1 (O'NEILL MICHAEL) 5 December 2002 (2002-12-05) paragraph '0020! - paragraph '0068! claims 1,12	30-34
X A	WO 01/16881 A (KONINKLIJKE PHILIPS ELECTRONICS N.V) 8 March 2001 (2001-03-08) page 3, line 1 - page 15, line 12; figures 1,2	1,4  2,3,5-34
	----- -/-	



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

### \* Special categories of cited documents:

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*Z\* document member of the same patent family

Date of the actual completion of the international search

24 June 2005

Date of mailing of the international search report

06/07/2005

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Schenkels, P

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/EP2004/000157

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 00/07113 A (CET TECHNOLOGIES PTE LTD; NG, YEW, LIAM; NG, KIM, CHWEE) 10 February 2000 (2000-02-10) page 2, line 20 - page 4, line 27 page 10, line 12 - page 29, line 4; figures 1-7 -----	1-30
A	PECK C C ET AL: "Genetic algorithm based input selection for a neural network function approximator with applications to SSME health monitoring" PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON NEURAL NETWORKS (ICNN). SAN FRANCISCO, MAR. 28 - APR. 1, 1993, NEW YORK, IEEE, US, vol. VOL. 1, 28 March 1993 (1993-03-28), pages 1115-1122, XP010111746 ISBN: 0-7803-0999-5 the whole document -----	1,30

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP2004/000157

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2002059154 A1	16-05-2002	NONE	
US 2002184569 A1	05-12-2002	NONE	
WO 0116881 A	08-03-2001	US 2002174079 A1 WO 0116881 A2 EP 1222628 A2 JP 2003508835 T	21-11-2002 08-03-2001 17-07-2002 04-03-2003
WO 0007113 A	10-02-2000	AU 4950899 A CN 1311880 A WO 0007113 A1 US 6470261 B1	21-02-2000 05-09-2001 10-02-2000 22-10-2002